

09748127.123700

```

1   TTTCTTCTGT TTGCTTACTC CCTATCCGGG GGCCCAAGGC GCTGTCTCCG
51  CCGCCCAAGC CCCGCGTAAA CCTGGGTGAC CTCGGAGACA TCCGTTGGAG
101 CATGAGTTCC CGACATCAGG CGGCGGCGGT GGTCCGGGAG AAACCCGGCG
151 GCGGGGAGAT AAGCCTGCCC AGGAGGCAGG GGGCTGGGCT AGCTGCCCCG
201 CCCC GCGCCT GACTTCGTTG GGGAGGGAGA CGCCCGGCTC CCGCCCCTAA
251 CTAGCCAGC CGCGCGGAGC GCCTGGGAGA GGAGAAGGAG CCGACCTGCC
301 GAGATGGAGG CGACCGGCAC CTGGGCGCTG CTGCTGGCGC TGGCGCTGCT
351 CCTGTGTGTG ACGCTGGCGC TGTCCGGGAC CAGGGCCCCG GGCACCTGC
401 CCCC GCGCCT CACGCCGCTA CCACTGCTGG GAAACCTCCT GCAGCTACGG
451 CCGGGGGCGC TGTATTACAG GCTCATGCGG CTGAGTAAGA AGTACGGACC
501 GGTGTTTACC ATCTACCTGG GACCTGGCG GCCTGTGGTG GTCCTGGTTG
551 GGCAGGAGGC TGTGCGGGAG GCCCTGGGAG GTCAGGCTGA GGAGTTCAGC
601 GCGCGGGGAA CCGTAGCGAT GCTGGAAGGG ACTTTTGATG GCCATGGGGT
651 TTTCTTCTCC AACGGGGAGC GGTGGAGGCA GCTGAGGAAG TTTACCATGC
701 TTGCTCTGCG GGACCTGGGC ATGGGGAAGC GAGAAGGCGA GGAGCTGATC
751 CAGGCGGAGG CCCGGTGTCT GGTGGAGACA TTCCAGGGGA CAGAAGGACG
801 CCCATTTCAT CCTCCCTGCG TGCTGGCCCA GGCCACCTCC AACGTAGTCT
851 GCTCCCTCCT CTTTGGCCTC CGCTTCTCCT ATGAGGATAA GGAGTTCAG
901 GCCGTGGTCC GGGCAGCTGG TGGTACCTCG CTGGGAGTCA GCTCCAGGG
951 GGGTCAGACC TACGAGATGT TCTCTGGTT CCTGCGGCC CTGCCAGGCC
1001 CCCACAAGCA GCTCCTCCAC CACGTCAGCA CCTTGGCTGC CTTACAGTC
1051 CGGCAGGTGC AGCAGCACCA GGGGAACCTG GATGCTTCGG GCCCGCACG
1101 TGACCTTGTC GATGCCCTCC TGCTGAAGAT GGCACAGGAG GAACAAAACC
1151 CAGGCACAGA ATTACCAAC AAGAACATGC TGATGACAGT CATTATTG
1201 CTGTTTGTCT GGACGATGAC GGTACGACC ACGGTCGGCT ATACCCTCCT
1251 GCTCCTGATG AAATACCTC ATGTCCAAA GTGGGTACGT GAGGAGCTGA
1301 ATCGGGAGCT GGGGGCTGGC CAGGCACCAA GCCTAGGGA CCGTACCCGC
1351 CTCCCTTACA CCGACGCGGT TCTGCATGAG GCGCAGCGGC TGCTGGCGCT
1401 GGTGCCCATG GGAATACCCC GCACCTCAT GCGGACCACC CGCTTCCGAG
1451 GGTACACCTT GCGCCAGGGC ACGGAGGTCT TCCCCTCCT TGGCTCCATC
1501 CTGCATGACC CCAACATCTT CAAGCACCCA GAAGAGTTCA ACCCAGACCG
1551 TTTCTGTGAT GCAGATGGAC GGTTCAGGAA GCATGAGGCG TTCCTGCCCT
1601 TCTCCTTAGG GAAGCGTGTG TGCCTTGGAG AGGGCCTGGC AAAAGCGGAG
1651 CTCTTCTCTT TCTTCAACCAC CATCTACAA GCCTCAAGCC CACCGTCAGT GGCCTTTTCA
1701 GTGCCCCCGG GACACCCTGA GCCTCAAGCC CACCGTCAGT GGCCTTTTCA
1751 ACATTTCCCC AGCCTTCCAG CTGCAAGTCC GTCCCACTGA CCTTCACTCC
1801 ACCACGCAGA CCAGATGAAG GAAGGCAACT TGGAAAGTGGT GGGTGCCGAG
1851 GACGGTGCCT CCAGCCTCAA CAGTGGGCAT GGACAGGGTT AATGCTTCCA
1901 GAGTGTACAC TGCAGGCAGC CACATTTACA CGCCTGCAGT TGTTCCTCGG
1951 AGTCTGTCCC ACGGCCACA CGCTCACTTG ACTCATGCTG CTAAGATGCA
2001 CAACCGCACA CCCATACACA ACTACAAGGG CCACAAAGCA ACTGCTGGGT
2051 TAGCTTTCCA CAGACATAAA TATAGTCCAT CTGCAATCAC AAGCACATAG
2101 CCAGGTAACC CACCAACTCC CCTGGATCTG CAGCCCACAC GTGGGAGTCT
2151 GGCTGTACAC TTCACAAGCC ACAGAAACGG CCACACATGT TCACAGCTCA
2201 CACGCCCTCT CCATTTCATG AACTTCTCAG TGTCCTGTG CCTGGTGCCT
2251 GGCACAGGGA ACAGCATGCC CCCTCCGGGG TCATGCCACC CAGAGACTGT
2301 CGCTGTCTAT GGGCCCAACT CATGCTCCCT CTCTTGGCTA CACCACTCTC
2351 CCAGCCTGTG ACCACCGATG TCCACACACC CCCAACCCT TGTCACACA
2401 GCTACCCACG TACGACATCG TCCTGGCTCC CCAGAGTATC TTCCCACTGA
2451 GACACGCGC CCCCACAGAG GCACAGTCCC CAGCCACCTC TGCAACTGCA
2501 GCCCTCAGTG ACCCCTTTT AAGCACCTTG ATTCTACCAA ATGCAAAACAC
2551 ATCTGGGTCT GCGATTATGC ACAGAGACTT TGGACATACG AGGACCTCA
2601 GACCGGAGGA ACACCTGCCC AACCCCAACA CGTGCTTATG TAACCAGTG
2651 GAAAGCGGCC CCTGTGCCC CTCCACACAC ACATACACAC TCACTGATCT
2701 ACAGCCCCTG TTCGGCGTCA GAGTCCCCAC TAGACCCAGT GGAAGGGGTT
2751 AGAGACCAAG TAGGGGCCAG TTTCCAATTC ACCCTGTCAG GGAGTGAGCC
2801 GGATCTGACG TTCCTTGTA CTTAAGGGTC CGGCTTGGGA ATTAAAGTTT
2851 GTTCTGCGC TTAGCCTAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
2901 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAA

```

(SEQ ID NO: 1)

FEATURES:
 5'UTR: 1-303
 Start Codon: 304
 Stop Codon: 1815
 3'UTR: 1818

HOMOLOGOUS PROTEINS:

gi 117254 sp P24461 CPG1_RABIT CYTOCHROME P450 2G1 (CYP1IG1) (P...	516	e-145
gi 404777 gb AAA31432.1 (L10912) cytochrome P-450 2B-Bx [Oryct...	513	e-144
gi 479930 pir S35666 cytochrome P450 2B4 isoform Bx - rabbit >...	513	e-144
gi 117212 sp P00178 CPB4_RABIT CYTOCHROME P450 2B4 (CYP1IB4) (P...	508	e-143
gi 320075 pir S31277 cytochrome P450 2B4-B1 - rabbit >gi 21369...	508	e-143
gi 4731350 gb AAD28466.1 (AF128849) cytochrome P450 2B10 relat...	508	e-143
gi 3123191 sp P04167 CPB2_RAT CYTOCHROME P450 2B2 (CYP1IB2) (P4...	508	e-143
gi 117213 sp P12789 CPB5_RABIT CYTOCHROME P450 2B5 (CYP1IB5) (P...	507	e-142
gi 89973 pir A27717 cytochrome P450 2B5, hepatic (form HP1) - ...	507	e-142
gi 2144292 pir O4RTP2 cytochrome P450 2B2 - rat	506	e-142

Est:

gb BE148597 BE148597 MR0-HT0241-150500-010-b02 HT0241 Homo sapi...	1091	0.0
gb BF359243 BF359243 RC6-ET0072-150600-011-F01 ET0072 Homo sapi...	779	0.0
gb AW753778 AW753778 RC1-CT0286-301099-011-f08 CT0286 Homo sapi...	652	0.0

EXPRESSION INFORMATION FOR MODULATORY USE:

gb|BE148597| head_neck
 gb|BF359243| lung_tumor
 gb|AW753778| colon

Tissue expression:
 Human leukocyte

00748127 122700

1 MEATGTWALL LALALLLLLT LALSGTRARG HLPPGPTPLP LLGNLLQLRP
 51 GALYSGLMRL SKKYGPVFTI YLGPWRPVVV LVGQEAAREA LGGQAEFSG
 101 RGTVMAMLEGT FDGHGVFFSN GERWRQLRKF TMLALRDLGM GKREGEELIQ
 151 AEARCLVETF QGTEGRPFDP SLLLAQATSN VVCSLLFGLR FSYEDKEFQA
 201 VVRAAGGTLL GVSSQGGQTY EMFSWFLRPL PGPHKQLLHH VSTLAAFTVR
 251 QVQQHQGNLD ASGPARDLVD AFLKMAQEE QNPGTEFTNK NMLMTVIYLL
 301 FAGTMTVSTT VGYTLLLLMK YPHVQKWVRE ELNRELGAGQ APSLGDRTRL
 351 PYTDAVLHEA QRLALVPMG IPRTLMTTR FRGYTLPQGT EVFPLLSIL
 401 HDPNIFKHPE ENPDREFDA DGRFRKHEAF LPFSLGKRV LGEGLAKAEL
 451 FLFFTTILQA FSLESPCPD TSLKPTVSG LFNIPPAFQL QVRPTDLHST
 501 TQTR

(SEQ ID NO:2)

FEATURES:

Functional domains and key regions:

Prosite search results:

[1] PDOC00004 PS00004 CAMP_PHOSPHO_SITE

cAMP- and cGMP-dependent protein kinase phosphorylation site

128-131 RKFT

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE

Protein kinase C phosphorylation site

Number of matches: 6

1 61-63 SKK
 2 99-101 SGR
 3 248-250 TVR
 4 288-290 TNK
 5 378-380 TTR
 6 473-475 SLK

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE

Casein kinase II phosphorylation site

Number of matches: 3

1 119-122 SNGE
 2 192-195 SYED
 3 343-346 SLGD

[4] PDOC00008 PS00008 MYRISTYL

N-myristoylation site

Number of matches: 10

1 51-56 GALYSG
 2 109-114 GTFDGH
 3 115-120 GVFFSN
 4 188-193 GLRFSY
 5 207-212 GTLLGV
 6 257-262 GNLDAS
 7 284-289 GTEFTN
 8 339-344 GQAPSL
 9 370-375 GIPRTL
 10 444-449 GLAKAE

[5] PDOC00009 PS00009 AMIDATION

Amidation site

Number of matches: 2

1 140-143 MGKR
 2 435-438 LGKR

[6] PDOC00029 PS00029 LEUCINE_ZIPPER

Leucine zipper pattern

Number of matches: 2

1 32-53 LPPGPTPLPLGNLLQLRPGAL
 2 39-60 LPLGNLLQLRPGALYSGLMRL

[7] PDOC00081 PS00086 CYTOCHROME_P450
Cytochrome P450 cysteine heme-iron ligand signature

433-442 FSLGKRVCLG

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	5	25	2.305	Certain
2	64	84	0.783	Putative
3	170	190	1.041	Certain
4	292	312	2.031	Certain
5	448	468	1.133	Certain

09748127 123700

BLAST Alignment to Top Hit:

>gi|117254|sp|P24461|CPG1_RABIT CYTOCHROME P450 2G1 (CYPIIG1)
(P450-NMB) (OLFACTIVE)
pir||B31944 cytochrome P450 2G1 - rabbit
Length = 494

Score = 516 bits (1315), Expect = e-145
Identities = 248/491 (50%), Positives = 345/491 (69%), Gaps = 3/491 (0%)

Query: 1 MEATGTWALLLALAL-LLLLTLALSCTRARGHLPPGPTPLPLLGNNLLQLRPGALYSGLMR 59
ME G + + LAL LL+ +A + G LPPGPTP+P LGNLLQ+R A + ++
Sbjct: 1 MELGGAFTIFLALCFSCLLILIAWKRVQKPGRLPPGPTPIPLGNNLLQVRDQTFQSFLK 60

Query: 60 LSKKYGPVFTIYLGPPWRPVVVLVGQEAVERALGGQAEFSGRGTVAMLEGTDFDGHGVFFS 119
L +KYGPVFT+Y+GP RPVV+L G EAV+EAL +A+EFSGRG +A +E F GHGV +
Sbjct: 61 LREKYGPVFTVYMGPP-RPVVILCGHEAVKEALVDRADFEFSGRGELASVERNFQGHGVALA 119

Query: 120 NGERWRQLRKFTMLALRDLMGKREGEELIQAEARCLVETFGQTEGRPFDPSSLQAQTS 179
NGERWR LR+F++ LRD GMGKR EE IQ EA L+E F+ T+G P DP+ L++ S
Sbjct: 120 NGERWRILRRFSLTILRDFGMGKRSIEERIQEEAGYLLEEFRTKKGAPIDPTFFLSRTVS 179

Query: 180 NVVCSLLFGLRFSYEDKEFQAVVRAAGGTLVGSSQGGQTYEMFSWFLRPLPGPHKQLLH 239
NV+ S++FG RF YEDK+F +++R + + +S+ Q Y+M+S ++ LPG H ++ +
Sbjct: 180 NVISSVVFSGSRFDYEDKQFLSLLRMINESFIEMSTPWAQLYDMYSGVMQYLPGRHNRIYY 239

Query: 240 HVSTLAAFTVRVQQHQNLDASGPARDLVDAFLKMAQEEQNPGTEFTNKNMLMTVIYL 299
+ L F +V+ ++ +LD P RD +D FL+KM Q++ NP TEF KN+++T +.L
Sbjct: 240 LIEELKDFIAARVKVNEASLDQNP-RDFIDCFILKMHQDKNNPHTEFNKLVLTTLNL 298

Query: 300 LFAGTMTVSTTVGYTLLLLMKYPHVQKWVREELNRELGAGQAPSLGDRTRLPYTDAVLHE 359
FAGT TVS+T+ Y LL+MK+P VQ + EE+N+ +G + PS+ DR ++P+TDAV+HE
Sbjct: 299 FFAGTETVSSTLRYGFLIMKHPEVQTKIYEEINQVIGPHRIPSVDDRVMKMFDTDAVHE 358

Query: 360 AQRLLALVPMGIPRTLMRTTRFRGYTLPGQTEVFPLLGSLHDPNIFKHPEEFNPDRFLD 419
QRL +VPMG+P ++R T FRGY LP+GT+VFPLLGSL+L DP F HP++F P FLD
Sbjct: 359 IQRLTDIVPMGVPHNVIRDTFRGYLLPKGTDVFPLLGSLVKDPKYFCHPDDFYPQHFLD 418

Query: 420 ADGRFRKHEAFLPFSLGKRVCLGEGLAELFLFFTILQAFSLSPCPDPTLSLKPTVS 479
GRF+K+EAF+PFS GKR+CLGE +A+ ELFL+FT+ILQ FSL PP + + P +S
Sbjct: 419 EQGRFRKNEAFVPFSSGKRICLGEAMARMELFLYFTSILQNFSLHPLVPPVNIDITPKIS 478

Query: 480 GLFNIPPAFQL 490
G NIPP ++L
Sbjct: 479 GFGNIPPTYEL 489

Hammer search results (Pfam):

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
PF00067	Cytochrome P450	594.4	6.9e-175	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00067	1/1	33	493	1	497	594.4	6.9e-175

```

1  NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
51  NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
1901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNTGACA TGGAGACACC
1951 TTGGATCGAA GAGGTCACAG CACCCTCCTC TTTCTTCCTC CCTACCCCA
2001 GCTGAGTAAG AAGTACGGAC CGGTGTTTAC CATCTACCTG GGACCCTGGC
2051 GGCCCTGTGGT GGTCTTGGTT GGGCAGGAGG CTGTGCGGGA GGCCCTGGGA
2101 GGTCAGGCTG AGGAGTTTAC CGGCCGGGGA ACCGTAGCGA TGCTGGAAGG
2151 GACTTTTGAT GGCCATGGTA AGTCAAGGGC TGCTAGGCCC TCCGCTCACA
2201 GCCTGCCACC ACTTACTGGT GTGTGACCTT TGCACATGGC TTAGTCCCTC
2251 TGTGTCCTCA TCTGTCAAAT GGAGTGATAA CAGTGCCCAT CAGCCGGGTG
2301 CAGTGGCTAG TGCCGTGAAAT CCAACACTT TGGGAGGCGG AGGTGGGTGG
2351 ATCACTTGAG GTCAGGAGTT CGAGACCAGC CTGGCCAACA TGGTGAAACC
2401 CTGTCTCTAC TAAAAATATA AAAATTAGCT GGGCATGGTG GTGCGTACCT
2451 GTAATCCAG ATACTTGGGA GGTGAGGCA GGAGAATCGC TTGAACCCGG
2501 GAGGCAGAGT TTGCAGTGAA CCAAGACTGT GCCACTGCAC TCCAGTCTGG
2551 GCAACAGAGT GAGCCTCCAT CTCAAACAAA CAAACAAAAA GCAGTGCCCA
2601 TCATGTAGGA TTGAGTGATT GAGTGAGGAC TGAGCCTTGT GCAAAGTGAG
2651 CACTCACTAA TCACCAGGTT GTAGTATCAG TGATAACCAT CAATGATCCA
2701 GGTAAAGCCC TGAGGGTTCA GAAAGATGCC GGAGCGCTTT CAAGGTGCTG
2751 GGGATTGGTG GGCAAGCCCT CGAATAATAG AAACAGTTCT CTGTATTACA
2801 ACAGAAAGCA GGAGGCCCAT GCTGGGTGCT GCCAGGAAC TTAGTAGTAA
2851 TAAGACAGCA CCGGTGCTGC TTCCCAGCG CACCTAGGCC AGTGGGAAA
2901 CAGACTCACC ACACAGTCCC AGCCAGAGT GGTGAGGCC AAGATGGGA
2951 AGCAGCGGGA GAAAGGTCAG GGTGGGATGG GGAGGGGTCA GGGCAAGAGG
3001 GGTCAGGGCC AGGCTGAGGG AAGCCCTGGG ACTGTAGGAA TTTAGAGGAG
3051 GTACCTGACC CGGCATGTTT GGTGAGGGAG ATTCAGGAAG TCTTCTTGGA
3101 AGAGAGGCTG TCGGAGCTGA GACTCATAAG ATGAGTGGGG AGGGTGTTC
3151 AGGCAGAAAG ACCAGCACCT ACAAAGCAT GACTTTGAGA GAAGCATTCA
3201 TCCATTCAAC TGATGAATTT TCAGACTGGG CACGCTGGCT CATGCCTGTA
3251 ATCCAGCAC TTTGGAAGGC TGAATGGGGA GGATGACTTG AGCCTAGGCA
3301 TTTGTGACAA GCCTGGGCAA CATGGTGAGA CCCTGCCCTC ACAAACAAA
3351 CAAACAAACA AAAAAATCATT ATACCTGGTA CCATGGGTAC CAGGTACATA
3401 GAAATGACTC AGGCAGATAT GGTGTCCTCT CCTACTGTGG GAGAGGCGGG
3451 CTTATACTGC AGTAAGACAA TAGAGGGAGG GAATATAATC CTAATAATGAG
3501 AGGTACAGAT TTGAGAGCAA ACACAGGGCA CAGGCATATG TACGAGGGTA

```

FIGURE 3, page 1 of 10

```

3551 AAGAGGGAAT CAGGGAAGGC TTCTCAGAGA AGGTGACATT TAAGCCGGGA
3601 CATGAAGGAT GAACGAGTTA GTTCACCAAG GATGGGATGG AAAGGGGTGA
3651 GAGTGATGGA GGCAGAGGGA ACTGCAGGAT CATAGGCCA GACAGGGGAT
3701 CCTGACGCCC TTGAGGAAGT GAGAGAAGAC CAGCGCAGTC GTAGTGGGTT
3751 AAGTAACAAA GCTGAGAAGC CAGGGAATC CCTGGTCATG CAGGGCCTGT
3801 GAGTCACGTC AGAGTGTTTG GGCTTTTGT TTTCTTGGGA GCAGTCGATT
3851 TTAAGCAGGG AACAGCTGTA TTCAGAGTTG GGAAGATCCT GTGGTTGCTG
3901 CCTGAAGGGG ATGAAACTGG AGGCTAGGAG CCCAGGGTGA TAGGGAGGAT
3951 CCAGGGTGAT GGGGAGGCTG GGAGGTCCGC GGTGATGGAC CAGGGCTGGG
4001 GCCAGGGGAT GGGGAGGAAG GAGTAATTGG GAGAGGCCTG GGGCTCTGGC
4051 CGAGGAATGG ATGGTGGGCT GAAACAGGGA GAGGAGAGAT GCTTAGGCCA
4101 CTTTGGAAAC CAGTAGGGCA AGGACAGGAG ACACCCAAGG GGAAGTGCCC
4151 AAGAGACCAC GACAGGCTGG CATTTGGACAG GGAAGGTCTG TCTGGAGCAG
4201 GTGTCTTGGG TAAGGGAGGA AAATGGTGCA GTTCCATCCT CCTCCCTCTC
4251 TGTTCACCTT CTAACCTACA TGGGGCACAG GACCCAGTGG GACTCCATAA
4301 ATGATGGGAT GGGTGGATGG AAGGAAGGAA GGAGGAAACA ACTCTTCATT
4351 CATCCTGGTT ATTTACAGAA CAGGCCAGGT GCGGTGCTCA CGCTTGCCAT
4401 TCTAGCACTT TGGGAGGCTG AGGTGGGTGG ATTACCTCAG GTCAGGAGTT
4451 CAAGACCAGC CTAGACAACG TAGAGAAACC CCATCTCTAC TGAAGATATA
4501 AAATTAGCTG GCGGTAGTGG CATATGCCTG TAATCCCAGC TAGTCGGGAA
4551 GCTGAGGCAG GAGAATCGCT TGAACCCGAG AGGCAGAGGT TGCGGTGAGC
4601 TGAGATCGTG CCATTGCACT CCAGCCTGGG TGACAAAGCA AGACCTCGTC
4651 TCAATAATAA TAATAATTAC AAAACAGAAG GAGCCTGGGT CATCCCAGCT
4701 ACCTACTTTT CAGGAGAATG TACTCCCTTA CCCAAGGGCA AAGGATGGGA
4751 GAACCAAGTT GATTATGCAT TTATTGAGCA CCTACTGAGT CCTCATCCCT
4801 GGGCTAGGCT GGAATGGACT CAGATGGAGC CTGAAGAGTC CCCCTCAGGG
4851 AACCTCACTA GAAAGAAGGA GGAATCGGCC GGGCGCGGTG GCTCACGCCCT
4901 GTAATCCCAA CACTTTGGGA GGCTGAGGTG GGTGGATCAC AAGGTCAGGA
4951 GATCGAGACC ATCCTGGCTA ACACAGTGAA ACCCCATCTC TACTAAAAAT
5001 ACAAAAAATG AGCCAGGCAT GGTGGCGGGC GCCTGTAGTC CCAGCTACTC
5051 AGGAGGCTGA GGCAGGAGAA TTGCTTGAAC CCGGGAGGCA GAGGTTGCAG
5101 TGAGACGAGA TCACGCCACT GCACTCCAGC CTGGGCAACA GAGCGAGATT
5151 CCGTCTCAAA AAAAAAAGA AAGAAAGGAA GAAGGGGAA TGGGGAGAG
5201 GGGCCGGTCC CTTTTGAGT CTAGCCTTCT GCGCAGGGGT TTTCTTCTCC
5251 AACGGGGAGC GGTGGAGGCA GCTGAGGAAG TTTACCATGC TTGCTCTGCG
5301 GGACCTGGGC ATGGGGAAGC GAGAAGGCGA GGAGCTGATC CAGGCGGAGG
5351 CCCGGTGTCT GGTGGAGACA TTCCAGGGGA CAGAAGGTCA GCATGGCGGG
5401 GTCACCCAGG GGTCTCCAGC CGAGTGAAAG GGAAACTCTC CTTACTGTGG
5451 CTGGGGGTGG CCCCACCCA GGTCTGGAA TGGGCAGGAG GGAAGCCCTT
5501 GAACTCTAGG GCTGGCCTGG GGGTTCTGTT CACTGCCACC TTCTGTCTCT
5551 GTCCCACTGT CTCTCCGAGG CTGTCTGAC ATCTCTCTGT GTGTCTCTGG
5601 TGCTATCATC CCATCTCTCC TGGGTCTCCA TCTCTCTCTC TGCTCTTTT
5651 CTTTCTCTCT CTTTCTCTCT ATTTTTTGGG CCCTCAGTCT ATCTCTGTTT
5701 CTGTCTCCCT GTCTGTGTGA TGGTCACTCT GTTTCTTTCT CCCTGTCTGT
5751 TTCTCTGTCC CTATCTGTCT GTATCCTTCT TTGCCTGTTT AGCTCTCTCC
5801 CTGCGCTGTC CATCATCTT TCCCTGCCTC CCGTCTCTCT TCTGGTTGGG
5851 TTCAGCCCCA ACCTGCTCCC CTCTGCCTGG CTCCATCACA GCCTACCTCC
5901 CTGCCCCCAT TCCCCCAGG ACGCCCATTC GATCCCCCCC TGCTGCTGGC
5951 CCAGGCCACC TCCAACGTAG TCTGCTCCCT CCTCTTTGGC CTCCGCTTCT
6001 CCTATGAGGA TAAGGAGTTC CAGGCCGTGG TCCGGGCAGC TGGTGGTACC
6051 CTGCTGGGAG TCAGCTCCCA GGGGGGTGAG GTGAGTGGGT GGGACCCCTC
6101 TCCAACCTACC TTCCCTGAAG GTTCTTGCCA AGGTCCCATG AGAACTAGCT
6151 GCCCTTCTCC CCACAGACCT ACGAGATGTT CTCCTGGTTC CTGCGGCCCC
6201 TGCCAGGCCC CCACAAGCAG CTCCTCCACC ACGTCAGCAC CTTGGCTGCC
6251 TTACACAGTCC GGCAGGTGCA GCAGCACCAG GGAACCTGG ATGCTTCGGG
6301 CCCCACAGT GACCTTGTCG ATGCCTTCCT GCTGAAGATG GCACAGGTGT
6351 GGAAGGGGTG CAGGGACCCC CTCTCTGAAT GGGCGTGGTG ACCTGGCAGG
6401 TCCCAGCCAG GTGTCCCTGG GGACCTCAAT TGGGTTCCCT TCTCTTTCTC
6451 TCTCTGCATG TCTCTGTGAG TATGAGTGTC TCTGTGCATG TGTGTGCATC
6501 CCTTCTCTGC ACATCTGTGC TGGCCCTTTC AGGGCGTTGC TCTCACTGCC
6551 TCTCCGCCCC CCGACCTGGG CATTGTGCCC GGGCTGTCTG TCTCTCCAGC
6601 ATCTCTCTCT TTTCTCCCTC CCACCTCGGC CCTGTGTGTT AGGCCCCATG
6651 CCCAGGGTCC TACACCAGCA ATCCCCAGGA TCACTTCATC CCATCCCTG
6701 CAGCCTCCCC AGACTTTTAT GTAAATTCAC AATTTTATGT GAATTATGGT
6751 CATTATTATG GAAGCCTTGC AATATCAAGT TATGTTAATA AAGTCCACTT
6801 TATTAATTAT ATAAGAACAA TATTTCTTTT CCTTTTTTTT TTCTTTTCTT
6851 TTTAAAGAGA CAGGATCTCT TTCTGTGCCC CAGGCTAGAG TACAGTTGCA
6901 AAATCATAGC TCACTGCAAC CTGGAACCTC TGGGCTCAAG CAATCCTCCT
6951 GCCTCGGGCT CCTGAGTAGC TGGGACAACA GGTGTGCACC ACCACACCTG
7001 GCTAAATTTT TTTTTTTTCT TTGTAGAGAT AGACTCTCAC TATGTTACCC
7051 AGGCTGTCT TGAATTCCTG GGCTCATGTA ATCCTCCTGC TGCCTTGAAC

```

09748137.122700

```
7101 TCCCAAAGTG CTGGGACTAT AGGCATAAGA CATCATGCCC GGTGCGGCAC
7151 AGTGGCTCAT GCCTGTAATC TCAGGACTTT GGGAGGCCGA GACGGGCGGA
7201 TCACCTGAGG TCGGGAGTTC GAGACCAGCC TGACCAACAT GGAGAAACCC
7251 CATCTCTACT AAAAAAATAA ATACAAAATT AGCCGGACGT GGTGGGCACAT
7301 GCCTGTAATC CCAGCTACTA GGGAGGCTGA GGCAGGAGAA TCGCTTGAAC
7351 CCGGAGAGCT TAGGTGCGG TGAGCTGAGA TTGCACCATT GCACTCCAGC
7401 CTGGGCAACA AGAGCGAAAT TCCATCTCAA AAAAAAATAA AAAGAAAAAA
7451 AGAAAAAAGA CACCATGCCC TATAAGTAAA CTAGAATTAA GGTGACTCCT
7501 AAGGAAATAA ATAGTTTTTA ACTGTACGAA CTTTGGGAAG AATGGGCGCA
7551 ATTCTTTAAT TAAATGCAGC CTCCTGTTT GTGGAGAAAG AAAAATTTTT
7601 CTTAACCCCTA TTGCCCCATT TCTTTCTCT TTTATTGAAT ATTTTTTAGT
7651 TTTAACTATA GTAAAATACA CATAACGTTT ACCATCTTAA CCATTTTTAG
7701 GTATACAGTA CAGTAGTGTT CAGTACATTC ATACTGTTAT GCAATCAGTC
7751 TCCAGAACTC TTCATGTTGC AAAGCTGAAA CTCTATACCC ATTAAACAAC
7801 TGCCTGTTCC TCCCTCCTCC AACCCTGGC AATCACCTT TTTTTTTGA
7851 GACGAAGTCT CACTCTGTCA CCCAGGCTAG AGTGCGGTGG CTCGATCTCG
7901 GCTCACTGCA AGCTCCGCCT CCCGGGTTCA TGCCATTCTC CTCCTCAGT
7951 CTCCTCAAGCA GCTGGGACTA CAGGTGCCCC TCACCACGCC TGGCTAATTT
8001 TTTGTATTTT TAGTAGAGAT GGAGTTTCAT CGTGTAGCC AGGCTGATCT
8051 CAAACTCCTG GCCTCAAGTG ATCCACCCGC CTCGGCCTCC CAAAGTGCTG
8101 GGACTACAGG CGTGAGCCAC TGTGCTGGC CAGGAAGTAG ACTCTTGATA
8151 TTAGTTCTCT CTGGTTGAAA TGTTTTTAAA AATGAAAGAG AATGACTAAT
8201 AACAAAAACA CAGAAAGTTA TAAGGATTGA TGAAGATGTG GAGACTTTGA
8251 AACCCTGTA TACCATTGGT GGGAAATGTG AACGACGCAG CCTGTGGAA
8301 AATGGTACAG CAGTTACCTG AGGTCAGGAG TTTGAAACCA ACCTGGCCAA
8351 CATGCAGAAA CCCCGTCTCC ATTAAATGTA CAAAAATTAG CCAGGNNNNN
8401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
10001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
10051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
10101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
10151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
10201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
10251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
10301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
10351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
10401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
10451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
10501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
10551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
10601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
```

FIGURE 3, page 3 of 10

10651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
10951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
11951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
12601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNCTCT	TTCTTCTTT	CTTCTCTCT
12651	CCCTTCTCTC	CTTCTCTCTC	TCCTGTCTTC	CTCTCTTTCT	CTCTTTCTTT
12701	CTTACAGGG	TCCTCTTTTG	TCTCCCAGGC	TGGGGTGCAG	TGGTACAAGC
12751	ATAGCTCACA	GCAGCCTTGA	ACTCCTAGGC	TCAAGTGATC	CTCCCACGTC
12801	AGCCTCCTGA	GCAGCTGGGA	CAACGGGCTC	ATACCACCAT	GCCTGGCTAA
12851	TTTTTTAATT	TTTCGTAGAG	ACAAGGTCTT	GTTATATTGC	CCAGGCTGGT
12901	CTCAACTTAT	TGGGCTCAAA	TGCTTCTCTC	ACCTCAGCCT	CCCACGTGGC
12951	TGGGATTACA	GGCATGAGCC	ACTGCACGCC	ACTCAACACT	CCACAAATGT
13001	TGATGCCATT	ATGTTTTGTG	AACTAGTGTC	CCTGGCACCC	GAGACTTGTA
13051	CTCCACACTC	GAGGACCAAA	TAGACTGGGG	TGGGAAGGGG	TTTATAGTTT
13101	CATTATTATT	TCCCTCAGG	GCACGGAGGT	CTTCCCCTC	CTTGCTCCA
13151	TCCTGCATGA	CCCCAACATC	TTCAAGCACC	CAGAAGAGTT	CAACCCAGAC
13201	CGTTTCCTGG	ATGCAGATGG	ACGGTTCAGG	AAGCATGAGG	CGTTCTTGCC
13251	CTTCTCCTTA	GGTATCTGCT	GCAGCCCTGG	GTATCACAAG	CAGGTGCTGG
13301	CGAACTCCAG	GCATCTGTGC	CAGCTGGGGG	CACCTTCTG	CACCTGGGG
13351	TTACTGTTGG	CTCCTCCACC	TGCTGTTCCC	CCCGTGGGCC	TGGGTGTGAG
13401	GAATACTGAC	TCAGCCCTCT	CTCTCTCTCT	CTCCTACCA	GGGAAGCGTG
13451	TCTGCCTTGG	AGAGGGCCTG	GCAAAAGCGG	AGCTCTTCTT	CTTCTTCAAC
13501	ACCATCCTAC	AAGCCTTCTC	CCTGGAGAGC	CGTGCCCGC	CGGTACACCC
13551	TGAGCCTCAA	GCCCACCGTC	AGTGGCCTTT	TCAACATTCC	CCCAGCCTTC
13601	CAGCTGCAAG	TCCGTCCCAC	TGACCTTCAC	TCCACCACGC	AGACCAGATG
13651	AAGGAAGGCA	ACTTGGAAGT	GGTGGGTGCC	CAGGACGGTG	CCTCCAGCCT
13701	CAACAGTGGG	CATGGACAGG	GTTAATGTCT	CCAGAGTGTA	CACTGCAGGC
13751	AGCCACATTT	ACACGCCTGC	AGTTGTTTTC	CGGAGTCTGT	CCCACGGCCC
13801	ACACGCTCAC	TTGACTCATG	CTGCTAAGAT	GCACAACCGC	ACACCCATAC
13851	ACAACACAAA	GGGCCACAAA	GCAACTGCTG	GGTTAGCTTT	CCACAGACAT
13901	AAATATAGTC	CATCTGCAAT	CACAAGCACA	TAGCCAGGTA	ACCCACCAAC
13951	TCCCTTGGAT	CTGCAGCCCA	CACGTGGGAG	TCTGGCTGTC	ACCTTCACAA
14001	GCCACAGAAA	CGGCCACACA	TGTTACAGC	TCACACGCCC	TCTCCATTCA
14051	TCGAACCTCT	CAGTGTCCCT	GTCCTGGTG	CCTGGCACAG	GGAACAGCAT
14101	GCCCCCTCCG	GGGTCAATGCC	ACCCAGAGAC	TGTCGCTGTC	TATGGCCCCA
14151	ACTCATGCTC	CCTCTCTTGG	CTACACCACT	CTCCCAGCCT	GTGACCACCG

FIGURE 3, page 4 of 10

14201 ATGTCCACAC ACCCCCAACC ACTTGTCAC ACAGCTACCC ACGTACGACA
14251 TCGTCTGGC TCCCCAGAGT ATCTTCCCAC TGAGACACGC CGCCCCCACA
14301 GAGGCACAGT CCCCAGCCAC CTCTGCAACT GCAGCCCTCA GTCACCCCTT
14351 TTTAAGCACC CTGATTCTAC CAAATGCAAA CACATCTGGG TCTGCGATTA
14401 TGCACAGAGA CTTTGGACAT ACGAGGACCC TCAGACCGGA GGAACACCTG
14451 CCCAACCCCA ACACGTGCTT ATGTAACCAC GTGGAAGCG GCCCTGCTG
14501 CCCCTCCACA CACACATACA CACTCACTGA TCTACAGCCC CTGTTGCGCG
14551 TCAGAGTCCC CACTAGACCC AGTGAAGGG GTTAGAGACC AAGTAGGGGC
14601 CAGTTTCCAA TTCACCCTGT CAGGGAGTGA GCCGGATCTG ACGTTCCCTG
14651 TGACTTAAGG GTCCGGCTTG GGAATTAAG TTTGTTTCTG GCCTTTAGCC
14701 TACTGCGTGT GTGACCCGTG TCAGTCACTG TGAGTAAGGG GTGGGGACAG
14751 GGGAGTCCAC CCCTCCCTTG AGGCTGGGCG GGAGCTGAAA AACATGGCCA
14801 CCGCCCACCC TGGCTGTTGA CATCAGGACC AGATGTGGAG CTGGGAGGAG
14851 GGGCAGGGCT GGTGACGCCC TGGGCTCAT TTCCAAAAAG GGCCAAGGTG
14901 TCCGGCGGTG GGAAGTGGGC AAGGAGGGGG TAACCCAAGC TGGACTGTGG
14951 ACCTTGGGGG CTTCCTCAGC CAGGAGAGC CTGAAGCCAA CTAGATCCAG
15001 ACCCTAGAGA CTCTTCAAAC TTGAGTACAG GAACTAGCTT GCAACACAGA
15051 CTCTAAGCCC ACTCCATT CTTCACCCT TTTTCTCTTG CCTCCCCTTC
15101 ACAAGGAAAC CAGAGGCATT TGTAATTTT CTTTCTTTT TTTTTTTTTT
15151 TTTTTTTTGA GACGGAGTCT CACTCTGTCA CCCAGGCTGG AGTGCAGTGG
15201 TGTGATCTG GCTCACTGCA GCCTCCGCT CCGGTTCAA GCCATTCTCC
15251 TGCTCAGCC TCCCAAGTAG CTGGGATTAC AGGTGTGTGC CACCACGCC
15301 AGCTAATTTT TGTATTTTGA GTAGAGATGG GGTTCACCA TGTGGCCAG
15351 GCTGGTCTCG AACTCCTGAC CTCAGATGAT CTGCCAGTCT CGGCCTCCCA
15401 AAATGCTGGG ATTACAGGCG TGAGTCGCTA CTAGATAAAT TTCTTATCTA
15451 GCAAAAGAG TTGCAAAACAT ACGCAAAAGT AGAAAGATAC AATGAGCCCC
15501 CAGGTGCCCA TCACCCAGCC TCATTTCAAT AGTCATCAAC TTTCTGCAGC
15551 TTTTACTTCA TCTATATCCT TTTCTGCCTC TTTTTTTTTT TTTTATTTT
15601 GAGATAGGGT TTTGCTTTGT TGCCCAAGCT GGGGTGCAGT AGCATGATCT
15651 CATAGTTTCA TGGGGCTTCA GACTCCTAGG CTCAGTGAT CCTCCCGCCT
15701 CGGCCTCCAA GCAGCTGGGA CTACAGATGC GTGCCACCAC ACCCAGCTAA
15751 ATTTCTTATT TTTATTTTCT ATAGAGAAAG TCTCACTATA CAGCCCTGTG
15801 CTGGTCTCAA ATTCCAGGCC TCAAGAGTTT CCATCCCAGC CTCCCAAAGT
15851 GCTGGGATTA TAGGCGTGAG TCACTGCACC CTGCCCTAAT ATTTTATTT
15901 TATCTATTGC TTTTATTTA CTTATTTATC TTTTATTTT GAGACAGAGT
15951 CTCCTCTGT GGGCCATGCT GGAGTGCAGT GGCATCATCT CGGCTCACTG
16001 TAACCTCCGC CTCTTAGGTT CAAGCAGTTC TCCTGCCTTG ACCTCCCGAG
16051 TAGCTGGAAT TACAGGTGCC TGCCACCAAG CCTGGCTAAT TTTTTTATTT
16101 GTAGTAGAGA TGGGGTTTTG CCATGTTGAC CAGGCTGGTC TCGAACTCCT
16151 GACCTCAGGT GATCTGCCCA CCTTGGCCTC CCAAAGTGCT GAGATTACTG
16201 GTATGAGCCA CCGTGCCTGG CCACCTATTG CTTTTTAAAG ATTATTTTTT
16251 TATTATTATT ATTTTATTTT TTGCAGATGG AGTTTCGCTA TTGTTGCCCA
16301 GGCTGGAGTG CAATGGCGTG ATCTCAGCTC ACCGCAACCT CCGCCTCCCA
16351 GGTTCAGCG ATTCTCCTGC CTCAGCATCC CTAGTAGCTG GGATTACAGG
16401 CATGCACCAC CATGTCCAGC TAATTTTGTA TTTTGTAGT AGACGAGGTT
16451 TCTCCAGGTT GGTGAGGCTG GTCTCAAAC CCCAACCTCA GGTGATCCGC
16501 CCATCTCGGC CTCCCAAAGT GCTGGGATTA CAGGTGTGAG CCACCGCGCC
16551 TGGCCTTAA GATTATTTTA AGGCAAATTA CAGAAAGCAA TTTAATGCAC
16601 ATTTCTGAGA GTTAAAGATA TTTTGGCCT TGACATTTTA TGAGGACAGT
16651 TTTCAAACAT GCAGCAAAGT TGAGGGAATT GTACAAGGAA CACCTTGTGC
16701 ACTCCTGCCT CAGTCTCCCA AGCAGCTGGG ACTACAGGTG CCCGTCACCA
16751 CGCCTGGCTA ATTTTTTGTA TTTTGTAGT AGATGGAGTT TCATCGTGT
16801 AGCCAGGCTG ATCTCAAAC CTGGCCTCA AGTGATCCAC CCGCCTCGGC
16851 CTCCCAAGT CTGGGACTAC AGGCGTGAGC CACTGTGCTT GGCCAGGTAA
16901 GTAGACTCTT GATATTAGTT CTCTCTGGTT GAAATGTTTT TAAAAATGAA
16951 AGAGAATGAC TAATAACAAA AACACAGAAA GTTATAAGGA TTGATGAAGA
17001 TGTGGAGACT TTGAAACCCA TGTATACCAT TGGTGGGAAT GTGAAACGAC
17051 GCAGCCCTGT GGAATATGGT ACAGCAGTTA CCTGAGGTCA GGAGTTTGAA
17101 ACCAACCTGG CCAACATGCA GAAACCCCGT CTCCATTAAA TGTACAAAAA
17151 TTAGCCAGGC ATGGTGGTGC GCACCTGTAA TCCCAGTAC TCGGGAGGCT
17201 GAGGCAGGAG AATTGCTTGA ACCCAGGAGG CGGAGGTTGC AGTGAGCCGA
17251 GATCGTGCCA CTGCACTCAG CCTGGGCAAC AAAGCAAGAC TCTGTCTCAA
17301 AAAAAAAAAG TCTACTTCCC AACCTTCCCA AAAATTTATC TAAACCCCGT
17351 GACAAAACCT TAACCTGTGT TTCCGACCCC AGGCTTGGCT GTTCTGGACA
17401 TTTACTTCCC AAAGGCTGTG TTCTCTCAGC CCCTCTGCCT GGTTTCTTTC
17451 AGGAGGAACA AAACCCAGGC ACAGAATTCA CCAACAAGAA CATGCTGATG
17501 ACAGTCATTT ATTTGCTGTT TGCTGGGACG ATGACGGTCA GCACCACGGT
17551 CGGCTATACC CTCCTGCTCC TGATGAAATA CCCTCATGTC CAAAGTAAGA
17601 GCCTTTTCCA CTTGCCAGGC CTTGGGAACA GAAGTCAGGG TTCTAGGCTG
17651 AGCAAGGTGG CTCACGCCTA TAATCCCAGC ACTTTGGGAG GCTGAGGCGG
17701 GCTGATCACT TGAGAATAGG AGTTTAAGAC CAGCCGGCCA ACACAGTGAA

FIGURE 3, page 5 of 10

17751 AC (SEQ ID NO: 3)

FEATURES:

Start:	1999	6349
Exon:	1999	2167
Intron:	2168	5236
Exon:	5237	5386
Intron:	5387	5919
Exon:	5920	6080
Intron:	6081	6166
Exon:	6167	6349
Intron:	6354	13542
Exon:	6354	6636
Intron:	6637	13087
Exon:	13088	13261
Intron:	13262	13441
Exon:	13442	13542
Intron:	13547	13648
Exon:	13547	13648
Stop:	13649	

SNPs:

DNA

Position	Major	Minor
----------	-------	-------

2226	G	C A
2226	C	A G T
3081	A	T G C
3788	A	T
3979	T	C G
5056	T	C G
5213	T	G A C
5508	A	C T G
5857	C	G A T
6385	C	T G
6813	C	T A
7853	C	T
12973	G	C T
12973	A	G T
13012	G	A C T
13072	A	T C G
13370	C	G
13682	C	T A
14631	C	T G

Context:

DNA

Position

2226	TGACAGGGGCCATGATGGAGACACCTTGGATCGAAGAGGTCACAGCACCCCTCCTCTTTCT TCCTCCCTACCCCCAGCTGAGTAAGAAGTACGGACCGGTGTTACCATCTACCTGGGACC CTGGCGGCCTGTGGTGGTCTGTTGGGCAGGAGGCTGTGCGGGAGGCCCTGGGAGGTCA GGCTGAGGAGTTGAGCGGCCGGGAACCGTAGCGATGCTGGAAGGGACTTTTGATGGCCA TGGTAAGTCAAGGGCTGCTAGGCCCTCCGCTCACAGCCTGCCACCACTTACTGGTGTGTG [G, C, A] CCTTTGCACATGGCTTAGTCCCTCTGTTGCCTCATCTGTCAAATGGAGTGATAACAGTGC CCATCAGCCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGGCGGAGGTGG GTGGATCACTTGAGGTGAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTGTCT CTACTAAAAATATAAAAAATTAGCTGGGCATGGTGGTGGTACCTGTAATCCCAGATACTT GGGAGGTTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGATGTTGCAGTGAACCAAGA
2226	TGACAGGGGCCATGATGGAGACACCTTGGATCGAAGAGGTCACAGCACCCCTCCTCTTTCT TCCTCCCTACCCCCAGCTGAGTAAGAAGTACGGACCGGTGTTACCATCTACCTGGGACC CTGGCGGCCTGTGGTGGTCTGTTGGGCAGGAGGCTGTGCGGGAGGCCCTGGGAGGTCA GGCTGAGGAGTTGAGCGGCCGGGAACCGTAGCGATGCTGGAAGGGACTTTTGATGGCCA TGGTAAGTCAAGGGCTGCTAGGCCCTCCGCTCACAGCCTGCCACCACTTACTGGTGTGTG [C, A, GsT]

CCTTTGCACATGGCTTAGTCCCTCTGTTGCCTCATCTGTCAAATGGAGTGATAACAGTGC
 CCATCAGCCGGGTGCAGTGGCTAGTGCCTGAAATCCCAACACTTTGGGAGGCGGAGGTGG
 GTGGATCACTTGAGGTGAGGATTCGAGACCAGCTGGCCAACATGGTGAAACCTGTCT
 CTACTAAAAATATAAAATAGCTGGGCATGGTGGTGCTTACCTGTAATCCAGATACTT
 GGGAGGTTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGATGTTGCAGTGAACCAAGA
 3081 AAACAGTTCTCTGTATTACAACAGAAAGCAGGAGGCCCATGCTGGGTGCTGCCAGGA
 CAGTAGTAACATAAGACAGCACCGGTGCTGCTTCCCCAGCGCACCTAGGCCAGTGGGAAA
 CAGACTACCACACAGTCCCAGCCAGAGTGGTCAGGGCCAAGATGGGGAAGCACGGGGA
 GAAAGGTGAGGTTGGGATGGGAGGGGTGAGGGCAAGAGGGGTGAGGGCCAGGCTGAGGG
 AAGCCCTGGGACTGTAGGAATTTAGAGGAGGTACCTGACCCGGCATGTTTGGTGAGGGAG
 [A, T, G, C]
 TTCAGGAAGTCTTCTGGAAGAGAGGCTGTGCGAGCTGAGACTCATAAGATGAGTGGGGA
 GGGTGTTCAGGCGAGAAAGACCAGCACCTACAAAGCATGACTTTGAGAGAAGCATTTCAT
 CCATTCAACTGATGAATTTTCAGACTGGGCACGCTGGCTCATGCTGTAATCCAGCACT
 TTGGAAGGCTGAATGGGAGGATGACTTGAGCCTAGGCATTTGTGACAAGCCTGGGCAAC
 ATGGTGAGACCCTGCCTCCACAAAACAAACAAACAAACAAACAAATCATTATACCTGGTAC
 3788 ATCCTAAATGAGAGGTACAGATTTGAGAGCAAACACAGGGCACAGGCATATGTACGAGG
 GTAAAGAGGGAATCAGGGAAGGCTTCTCAGAGAAGGTGACATTTAAGCCGGGACATGAAG
 GATGAACGAGTTAGTTTACCAAGGATGGGATGGAAGGGGTGAGAGTGATGGAGGCAGAG
 GGAAGTGCAGGATCATAGGCTTAGACAGGGGATCCTGACGCCCTTGAGGAAGTGAGAGAA
 GACCAGCGCAGTCGATAGTGGGTTAAGTAACAAAGCTGAGAAGCCAGGGAATCCCTGGTC
 [A, T]
 TGCAGGGCCTGTGAGTCACGTCAGAGTGTTTGGGCTTTTGTCTTCTGGGAGCAGTCGA
 TTTTAAGCAGGGAACAGCTGTATTTCAGAGTTGGGAAGATCCTGTGGTTGCTGCCTGAAGG
 GGATGAAACTTGGAGGCTAGGAGCCAGGGTGATAGGGAGGATCCAGGGTGATGGGGAGGC
 TGGGAGGTCCGCGGTGATGGACAGGGCTGGGGCCAGGGGATGGGGAGGAAGGAGTAATT
 GGGAGAGGCTGGGCTCTGCGCCAGGAATGGATGGTGGGCTGAAACAGGGAGAGGAGAG
 3979 ATCATAGGCTTAGACAGGGGATCCTGACGCCCTTGAGGAAGTGAGAGAAGACCAGCGCAG
 TCGTAGTGGGTAAAGTAACAAAGCTGAGAAGCCAGGGAATCCCTGGTCATGCAGGGCCT
 GTGAGTCACGTACAGTGTTTGGGCTTTTGTCTTCTGGGAGCAGTCGATTTTAAGCAG
 GGAACAGCTGTATTTCAGAGTTGGGAAGATCCTGTGGTTGCTGCCTGAAGGGGATGAACT
 GGAGGCTAGGAGCCAGGGTGATAGGGAGGATCCAGGGTGATGGGGAGGCTGGGAGGTCC
 [T, C, G]
 CCGTGATGGACAGGGCTGGGGCCAGGGATGGGGAGGAAGGAGTAATTGGGAGAGGCCT
 GGGGCTCTGGCCAGGAATGGATGGTGGGCTGAAACAGGGAGAGGAGATGCTTAGGCC
 ACTTTGGAACACAGTAGGCAAGGACAGGAGACACCAAGGGGAAGTGCCCAAGAGACCA
 CGACAGGCTGGCATTGGACAGGGAAGGTCTGTCTGGAGCAGGTGCTTGATAGGGAGG
 AAAATGGTGACGTTCCATCCTCCTCCTCTCTGTTCAACCTCTAACTACATGGGGCACA
 5056 AGTTTGATTATGCATTTATTGAGCACCTACTGAGTCCTCATCCCTGGGCTAGGCTGGAAT
 GGACTCAGATGGAGCCTGAAGAGTCCCCCTCAGGGAACCTCACTAGAAAGAAGGAGGAAT
 CGGCCGGGCGCGGTGGCTCAGCCTGTAATCCCAACACTTTGGGAGGCTGAGGTGGGTGG
 ATCACAAGTTCAGGAGATCGAGACCATCCTGGCTAACACAGTGAACCCCATCTCTACTA
 AAAATACAAAAATGAGCCAGGCATGGTGGCGGGCGCTGTAGTCCCAGCTACTCAGGAG
 [T, C, G]
 CTGAGGCAGGAGAATTGCTTGAACCCGGGAGGCAGAGGTTGAGTGAGACGAGATCACGC
 CACTGCACCTCCAGCCTGGGCAACAGAGCGAGATTCCGTCTCAAAAAAAGAAAGAAA
 GGAAGAAGGGGAATGGGGAGAGGGGCGGTCCTTTTGTAGTCTAGCCTTCTGCGCAG
 GGGTTTCTTCTCAACGGGAGCGGTGGAGGCAGCTGAGGAAGTTTACCATGCTTGCTC
 TGCGGGACCTGGGCATGGGAAGCGAGAAGCGAGGAGCTGATCCAGCGGAGGCCCGGT
 5213 CTTTGGGAGGCTGAGGTGGGTGGATCACAAGGTCAGGAGATCGAGACCATCCTGGCTAAC
 ACAGTGAAACCCATCTCTACTAAAAATACAAAAATGAGCCAGGCATGGTGGCGGGCGC
 CTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGCTTGAACCCGGGAGGCAGA
 GGTTCAGTGAGAGAGATCAGCCACTGCACTCCAGCCTGGGCAACAGAGCGAGATTCC
 GTCTCAAAAAAAGAAAGAAAGAAAGGAAGAGGGGAATGGGGAGAGGGGCGGTCCT
 [T, G, A, C]
 TTTGAGTCTAGCCTTCTGCGCAGGGGTTTCTTCTTCAACGGGGAGCGGTGGAGGCAGCT
 GAGGAAGTTTACCATGCTTGCTCTGCGGGACCTGGGCATGGGGAAGCGAGAAGCGAGGA
 GCTGATCCAGCGGAGGCGCGGTGTCTGGTGAGACATTCAGGGGAGAGAAGGTGAGCA
 TGGCGGGTCAACCCAGGGTCTCCAGCCGAGTGAAGGGAAACTCTCTACTGTGGCTG
 GGGTGGCCCCAACCCAGGTCTTGAATGGGCAGGAGGGAAGCCTTGAACCTTAGGGCT
 5508 TCCCTTTTGTAGTCTAGCCTTCTGCGCAGGGGTTTCTTCTTCAACGGGGAGCGGTGGAG
 GCAGCTGAGGAAGTTTACCATGCTTGCTCTGCGGGACCTGGGCATGGGGAAGCGAGAAGG
 CGAGGAGCTGATCCAGGCGGAGGCCCGGTGTCTGGTGAGACATTCAGGGGACAGAAGG
 TCAGCATGGCGGGGTCAACCCAGGGTCTCCAGCCGAGTGAAGGGAAACTCTCTACTG
 TGGCTGGGGGTGGCCCCAACCCAGGTCTTGAATGGGCAGGAGGGAAGCCTTGAACCTT

FIGURE 3, page 7 of 10

[A, C, T, G]

GGGCTGGGCTGGGGGTTCTGTCTACTGCCACCTTCTGTCTCTGTCCACTGTCTCTCCGA
GGCTGTGATGACATCTCTGTGTGTCTCTGGTGCTATCATCCCATTTCTCCTGGGTCTC
CATCTCTCTCTGTCTCTTTTCTTTCTCTCTCTCTTTCTCTATTTTGGGCCCTCAGT
CTATCTCTGTTTCTGTCTCCCTGTCTGTGTGATGGTCACTCTGTTTCTTTCTCCCTGTCT
GTTTCTCTGTCCCTATCTGTCTGTATCTTCTTTGCCTGTTTAGCTCTCTCCCTGCGCTG

5857 CTGTCTCTCCGAGGCTGTGATGACATCTCTGTGTGTCTCTGGTGCTATCATCCCATTC
TTCTGGGTCTCCATCTCTCTCTGTCTCTTTTCTTTCTCTCTCTTTCTCTATTTT
TGGGCCCTCAGTCTATCTGTCTGTCTCCCTGTCTGTGTGATGGTCACTCTGTTTCT
TTCTCCCTGTCTGTTTCTCTGTCTCCCTATCTGTCTGTATCTCTTTGCTGTTTAGCTCT
CTCCCTGCGCTGTCCATCCATCTTTCCCTGCTCCCTGTCTCTCTCTGTTGGGTTTCAGC
[C, G, AsT]
CCAACCTGCTCCCTCTGCTGGCTCCATCACAGCCTACCTCCCTGCCCCCATTTCCCCC
AGGACGCCCATTCGATCCCTCCCTGCTGCTGGCCAGGCCACCTCCAACGTAGTCTGCTC
CCTCCTCTTTGGCCTCCGCTTCTCTATGAGGATAAGGAGTCCAGGCCGTGGTCCGGGC
AGCTGGTGGTACCTGCTGGGAGTCACTCCAGGGGGGTGAGGTGAGTGGGTGGGACCC
CTCTCCAACCTACCTTCCCTGAAGGTTCTTGCCAAGGTCCCATGAGAACTAGCTGCCCTTC

6385 GTGGGTGGGACCCCTCTCCAACCTACCTTCCCTGAAGGTTCTGCCAAGGTCCCATGAGAA
CTAGCTGCCCTTCTCCCCACAGACCTACGAGATGTTTCTCCTGGTTCCCTGCCGCCCTGCC
AGGCCCCACAAAGCAGCTCCTCCACCAGTCAACCTTGGCTGCCTTCACAGTCCGGCA
GGTGACAGCAGCACCAGGGGAACCTGGATGCTTCGGGCCCGCACGTGACCTTGTGATGC
CTTCTGCTGAAGATGGCACAGGTGTGGGAAGGTGACAGGACCCCTCTCTGAATGGGC
[C, T, G]
TGGTGACCTGGCAGGTCCAGCCAGGTGTCCCTGGGGACCTCAATTGGGTTCTCTCTCT
TTCTCTCTCTGATGTCTCTGTGAGTATGAGTGTCTCTGTGATGTGTGTGATCCCTTC
TCTGCACATCTGTCTGGCCCTTTCAGGGCGTTGCTCTCACTGCCTCTCCCGCCCCGAC
CTGGGCATTTGTGCGGGCTGTCTGTCTCTCCAGCATCTCTCTCTTTCTCCCTCCCAAC
TCGGCCCTTGTGTTTCAAGCCCATGCCAGGGTCTACACCAGCAATCCCAAGGATCACT

6813 ATCTGTGCTGGCCCTTTCAGGGCGTTGCTCTCACTGCCTCTCCCGCCCCGACCTGGGCA
TTTGTGCCGGCTGTCTGTCTCTCCAGCATCTCTCTCTTTCTCCCTCCCACTCGGCC
TTGTGTTTCAAGCCCATGCCAGGGTCTACACCAGCAATCCCAAGGATCACTTCATCCC
ATCCCTGACGCTCCCAAGCTTTTATGTAATTCACAATTTTATGTGAATTATGGTCA
TTTATTAGGAAGCCTTGCAATATCAAGTTATGTTAATAAGTCCACTTTATTAATTATAT
[C, T, A]
AGAACAATATTTCTTTCTTTTCTTTTCTTTTCTTTTAAAGAGACAGGATCTCTTTC
TGTTGCCAGGCTAGAGTACAGTTGCAAAATCATAGCTCACTGCAACCTTGAACCTCCTGG
GCTCAAGCAATCCTCCTGCCTCGGGCTCCTGAGTAGCTGGGACAACAGGTGTGACCAACC
ACACCTGGCTAAATTTTCTTTTCTTTTGTAGAGATAGACTCTCACTATGTTACCCAGG
CTGGTCTTTGAATCCTGGGCTCATGTAATCCTCCTGCTGCCTTGAACCTCCCAAGTGTCTG

7853 TCTTTAATTAAATGCAGCCTCCCTGTTTGTGGAGAAAGAAAAATTTTCTTAACCCATT
GCCCATTTCTTTCTCTTTTATGTAATATTTTATGTTTAACTATAGTAAATAACACA
TAACGTTTACCATCTTAACCATTTTATAGGTATACAGTACAGTAGTGTTCAGTACATTCAT
ACTGTTATGCAATCAGTCTCCAGAACTCTCATGTTGCAAGCTGAAACTCTATACCCAT
TAAACAATGCTGTTCTCCTCCTCCAAACCCCTGGCAATCACCTTTTTTTTTTTGAGA
[C, T]
GAAGTCTCACTCTGTACCCAGGCTAGAGTGGGTGGCTCGATCTCGGCTCACTGCAAGC
TCCGCTCCCGGGTTCATGCCATTCTCCTGCCTCAGTCTCCCAAGCAGCTGGGACTACAG
GTGCCGCTCACCACGCTGGCTAATTTTTTGTATTTTATAGAGATGGAGTTTCATCGT
GTTAGCCAGGCTGATCTCAAACTCCTGGCTCAAGTATCCACCGCTCGGCCTCCCAA
AGTGTGGGACTACAGGCTGAGCCACTGTGCCTGGCCAGGAAGTAGACTCTTGATATTA

12973 CTGTCTCTCTCTTTCTCTCTTTCTTTCTTTGACAGGGTCTCTCTTTGTCTCCAGGCTG
GGGTGCACTGGTACAAGCATAGCTCACAGCAGCTTGAACCTCCTAGGCTCAAGTGATCCT
CCCACGTGACCTCCTGAGCAGCTGGGACAACGGGCTCATACCACCATGCCTGGCTAATT
TTTTAATTTTCTGAGAGACAAGGTCTTGTATATGCCCAGGCTGGTCTCAAACTCCTG
GGCTCAAATGCTTCTCTACCTCAGCTCCACGCTGGCTGGGATTACAGGCATGAGCCAC
[G, C, T]
GCACGCCACTCAACACTCCACAAATGTTGATGCCATTATGTTTTGTGAACCTAGTGTCCCT
GGCACCAGAGACTTGTACTCCACTCGAGGACCAATAGACTGGGGTGGGAAGGGGTTT
ATAGTTTCATTATTTTCCCTCAGGGCACGGAGGTCTTCCCCCTCCTTGGCTCCATCC
TGCATGACCCCAACATCTTCAAGCACCAGAGAGTTCAACCCAGACCGTTTCTGGATG
CAGATGGACGGTTCAAGGAAGCATGAGGCGTTCTGCTTCTCTTAGGTATCTGCTGCA

12973 CTGTCTTCCTCTCTTTCTCTCTTTCTTTCTTGACAGGGTCTCTCTTTGTCTCCCAGGCTG
GGGTGCAGTGGTACAAGCATAGCTCACAGCAGCCTTGAACCTCTAGGCTCAAGTGATCCT
CCCACGTCAGCCTCCTGAGCAGCTGGGACAACGGGCTCATACCCATGCCTGGCTAATT
TTTTAATTTTTCTGAGAGACAAGGTCTTGTATATTGCCAGGCTGGTCTCAAACCTCTG
GGCTCAAATGCTTCTCTACCTCAGCCTCCACGTTGGTGGGATTACAGGCATGAGCCAC
[A, G, T]
GCACGCCACTCAACACTCCACAAATGTTGATGCCATTATGTTTGTGAACTAGTGCCCT
GGCACCCGAGACTTGTACTCCACACTCGAGGACCAAATAGACTGGGGTGGGAAGGGGTTT
ATAGTTTCATTATTTATTTCCCTCAGGGCACGGAGGTCTTCCCCCTCCTGGCTCCATCC
TGCATGACCCCAACATCTTCAAGCACCCAGAAGAGTTCAACCCAGACCGTTTCTTGGATG
CAGATGGACGGTTAGGAAGCATGAGGCGTTCTGCCCCTCTCCTTAGGTATCTGCTGCA

13012 CTCTCTTTGTCTCCCAGGCTGGGGTGCAGTGGTACAAGCATAGCTCACAGCAGCCTTGAA
CTCCTAGGCTCAAGTGATCCTCCACGTCAGCCTCCTGAGCAGCTGGGACAACGGGCTCA
TACCACCATGCTGGCTAATTTTTTAATTTTTCTGAGAGACAAGGTCTTGTATATTGCC
CAGGCTGGTCTCAAACCTCCTGGGCTCAAATGCTTCTCTCACCTCAGCCTCCACGTTGGCT
GGGATTACAGGCATGAGCCACTGCACGCCACTCAACACTCCACAAATGTTGATGCCATTA
[G, A, C, T]
GTTTGTGAACTAGTGCTCCTGGCACCCGAGACTTGTACTCCACACTCGAGGACCAAATA
GACTGGGGTGGGAAGGGGTTTATAGTTTCATTATTATTTCCCCCTCAGGGCACGGAGGTCT
TCCCCCTCCTTGGCTCCATCCTGCATGACCCCAACATCTTCAAGCACCCAGAAGAGTTCA
ACCCAGACCGTTTCTTGGATGCAGATGGACGGTTCAGGAAGCATGAGGCGTTCTGCCCC
TCTCCTTAGGTATCTGCTGCAGCCCTGGGTATCACAAGCAGGTGCTGGCGAATCCAGGC

13072 CTCCTAGGCTCAAGTGATCCTCCACGTCAGCCTCCTGAGCAGCTGGGACAACGGGCTCA
TACCACCATGCTGGCTAATTTTTTAATTTTTCTGAGAGACAAGGTCTTGTATATTGCC
CAGGCTGGTCTCAAACCTCCTGGGCTCAAATGCTTCTCTCACCTCAGCCTCCACGTTGGCT
GGGATTACAGGCATGAGCCACTGCACGCCACTCAACACTCCACAAATGTTGATGCCATTA
TGTTTTGTGAACTAGTGCTCCTGGCACCCGAGACTTGTACTCCACACTCGAGGACCAAAT
[A, T, C, G]
GACTGGGGTGGGAAGGGGTTTATAGTTTCATTATTATTTCCCCCTCAGGGCACGGAGGTCT
TCCCCCTCCTTGGCTCCATCCTGCATGACCCCAACATCTTCAAGCACCCAGAAGAGTTCA
ACCCAGACCGTTTCTTGGATGCAGATGGACGGTTCAGGAAGCATGAGGCGTTCTGCCCC
TCTCCTTAGGTATCTGCTGCAGCCCTGGGTATCACAAGCAGGTGCTGGCGAATCCAGGC
ATCTGTGCCAGCTGGGGGCACCTTCTGCACCTGGGCTTACTGTTGGCTCCTCCACCTG

13370 ATAGACTGGGGTGGGAAGGGGTTTATAGTTTCATTATTATTTCCCCCTCAGGGCACGGAGG
TCTTCCCCCTCCTTGGCTCCATCCTGCATGACCCCAACATCTTCAAGCACCCAGAAGAGT
TCAACCCAGACCGTTTCTTGGATGCAGATGGACGGTTCAGGAAGCATGAGGCGTTCTTGC
CCTTCTCCTTAGGTATCTGCTGCAGCCCTGGGTATCACAAGCAGGTGCTGGCGAATCCA
GGCATCTGTGCCAGCTGGGGGCACCTTCTGCACCTGGGCTTACTGTTGGCTCCTCCAC
[C, G]
TGCTGTTCCCCCGTGGGCTGGGTGTGAGGAATACTGACTCAGCCCTCTCTCTCTCTCT
CTCCTCACCAGGAAGCGTGTCTGCCTTGGAGAGGGCTGGCAAAAGCGGAGCTCTTCTCT
CTTCTTACCACCATCTTACAAGCCTTCTCCTGGAGAGCCCGTGCCCGCCGGTACACCC
TGAGCCTCAAGCCACCGTCAGTGGCCTTTTCAACATTCCCCCAGCCTTCCAGCTGCAAG
TCCGTCCTCCTGACCTTCACTCCACCACGACAGACAGATGAAGGAAGGCAACTTGAAGT

13682 CCGTGGGCTGGGTGTGAGGAATACTGACTCAGCCCTCTCTCTCTCTCTCTCCTCACCAG
GGAAGCGTGTCTGCCTTGGAGAGGGCTGGCAAAAGCGGAGCTCTTCTCTCTCTTACCA
CCATCCTACAAGCCTTCTCCTGGAGAGCCCGTGCCCGCCGGTACACCTGAGCCTCAAG
CCCACCGTCAGTGGCCTTTTCAACATTCCCCCAGCCTTCCAGCTGCAAGTCCGTCCACT
GACCTTCACTCCACCACGACAGACAGATGAAGGAAGGCAACTTGAAGTGGTGGGTGCC
[C, T, A]
GGACGGTGCCTCCAGCCTCAACAGTGGGCATGGACAGGGTTAATGTCTCCAGAGTGTA
CTGACGGCAGCCACATTTACACGCTGCAGTTGTTTTCCGGAGTCTGTCCACGGCCAC
AGCCTCACTTGAATCATGCTGCTAAGATGCACAACCGCACACCCATACACAATACAAGG
GCCACAAAGCAACTGCTGGGTTAGCTTTCCACAGACATAAATATAGTCCATCTGCAATCA
CAAGCACATAGCCAGGTAACCCACCAACTCCCCGATCTGCAGCCACACGTTGGGAGTC

Figure 1 consists of 14 bar charts, labeled (a) through (l), each representing a different demographic or attitudinal variable. The y-axis for all charts is 'Percentage of respondents' ranging from 0 to 100. The x-axis for each chart has four categories: 'No party affiliation', 'Party affiliation', 'Party identification', and 'Party loyalty'. The data is as follows:

- (a) Age: 18-24 (100%), 25-34 (100%), 35-44 (100%), 45-54 (100%), 55-64 (100%), 65-74 (100%), 75-84 (100%), 85-94 (100%), 95-104 (100%).
- (b) Sex: Male (100%), Female (100%).
- (c) Education: Less than high school (100%), High school (100%), Some college (100%), Bachelor's degree (100%), Master's degree (100%), Doctorate (100%).
- (d) Income: Less than \$10,000 (100%), \$10,000-\$19,999 (100%), \$20,000-\$29,999 (100%), \$30,000-\$39,999 (100%), \$40,000-\$49,999 (100%), \$50,000-\$59,999 (100%), \$60,000-\$69,999 (100%), \$70,000-\$79,999 (100%), \$80,000-\$89,999 (100%), \$90,000-\$99,999 (100%), \$100,000 or more (100%).
- (e) Employment: Not employed (100%), Employed (100%).
- (f) Religion: No religion (100%), Protestant (100%), Catholic (100%), Jewish (100%), Muslim (100%), Other (100%).
- (g) Political affiliation: No affiliation (100%), Democrat (100%), Republican (100%), Independent (100%), Other (100%).
- (h) Party affiliation: No party affiliation (100%), Democrat (100%), Republican (100%), Independent (100%), Other (100%).
- (i) Party identification: No party identification (100%), Democrat (100%), Republican (100%), Independent (100%), Other (100%).
- (j) Party loyalty: No party loyalty (100%), Democrat (100%), Republican (100%), Independent (100%), Other (100%).
- (k) Party support: No party support (100%), Democrat (100%), Republican (100%), Independent (100%), Other (100%).
- (l) Party preference: No party preference (100%), Democrat (100%), Republican (100%), Independent (100%), Other (100%).

ePCR to dbSTS

FIGURE 3, page 10 of 10